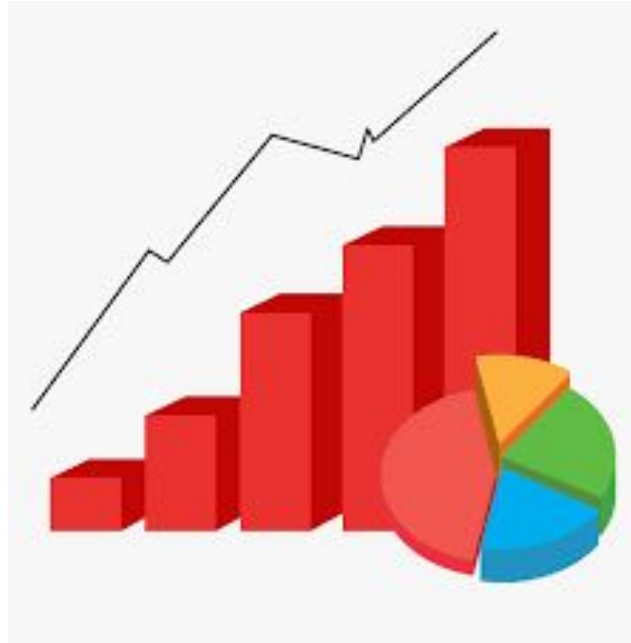# Statistics

Data

data (KS1): Information of a quantitative nature consisting of counts or measurements. Initially data are nearly always counts or things like percentages derived from counts. When they refer to measurements that are separate and can be counted, the data are discrete. When they refer to quantities such as length or capacity that are measured, the data are continuous. Singular: datum. database A means of storing sets of data.

continuous data (KS1): Data arising from measurements taken on a continuous variable (examples: lengths of caterpillars; weight of crisp packets). Continuous data may be grouped into touching but non-overlapping categories. (Example height of pupils [x cm] can be grouped into 130 ≤ x < 140; 140 ≤ x <150 etc.) Compare with discrete data.

discrete data (KS3): Data resulting from situations involving discrete variables Examples: value of coins in pupils' pockets; number of peas in a pod). Discrete data may be grouped. Example: Having collected the shoe sizes of pupils in the school, the data might be grouped into 'number of pupils with shoe sizes 3 – 5, 6 – 8, 9 – 11' etc.

grouped (discrete data) (KS3): Observed data arising from counts and grouped into non-overlapping intervals. Example: score in test / number of children obtaining the score scores 1 – 10, 11 – 20, 21 – 30, 31 – 40, 41 – 50 etc. In this example there are equal class intervals.

categorical data (KS1): Data arising from situations where categories (unordered discrete) are used. Examples: pets, pupils' favourite colours; states of matter – solids, liquids, gases, gels etc; nutrient groups in foods – carbohydrates, proteins, fats etc; settlement types – hamlet, village, town, city etc; and types of land use – offices, industry, shops, open space, residential etc

equally likely (KS3): In an experiment (trial in statistics) the result is the outcome. Two outcomes are equally likely if they have the same theoretical probability of occurrence. Example: when an unbiased coin is tossed the two outcomes 'head' or 'tail' are equally likely.

frequency

frequency (KS1): The number of times an event occurs; or the number of individuals (people, animals etc.) with some specific property. frequency density See histogram.

table (KS1): An orderly arrangement of information, numbers or letters usually in rows and columns. 2. See multiplication table

tally (KS1): Make marks to represent objects counted; usually by drawing vertical lines and crossing the fifth count with a horizontal or diagonal strike through. A Tally chart is a table representing a count using a Tally.

Favourite Pets		
Pet	**Tally Marks**	**Number**
Cat	⅄⅄ ⅄⅄	10
Dog	||||	4
Rabbit	⅄⅄ |	6

frequency table (KS3): A table for displaying a set of observations showing how frequently each event or quantity occurs in a statistical trial. This is an example of a frequency distribution, which sometimes can also be represented algebraically or graphically.

Charts and graphs
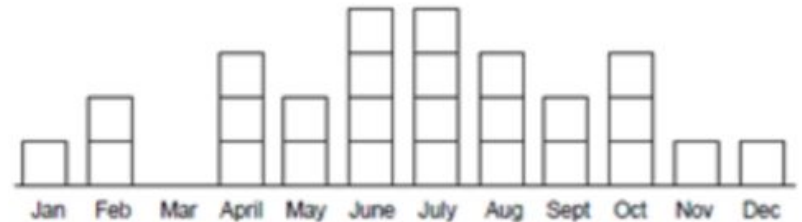
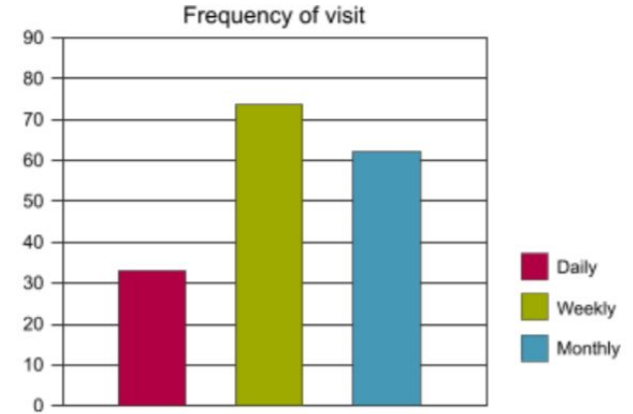**chart (KS1):** Another word for a table or graph

graph (KS2): A diagram showing a relationship between variables. Adjective: graphical.

bar chart (KS1): A format for representing statistical information. Bars, of equal width, represent frequencies and the lengths of the bars are proportional to the frequencies (and often equal to the frequencies). Sometimes called bar graph. The bars may be vertical or horizontal depending on the orientation of the chart.

Pictogram (KS1): A format for representing statistical information. Suitable pictures, symbols or icons are used to represent objects. For large numbers one symbol may represent a number of objects and a part symbol then represents a rough proportion of the number.

column graph (KS1): A bar graph where the bars are presented vertically.

block graph (KS1): A simple format for representing statistical information. One block represents one observation. Example: A birthday graph where each child places one block, or colours one square, to represent himself / herself in the month in which he or she was born.

Frequency of visit

Daily
Weekly
Monthly

Jan   Feb   Mar   April   May   June   July   Aug   Sept   Oct   Nov   Dec

Charts and graphs

Line graph: A graph in which adjacent points are joined by straight-line segments. Such a graph is better seen as giving a quick pictorial visualisation of variation between points rather than an accurate mathematical description of the variation between points.

Pie chart (KS2): Also known as pie graph. A form of presentation of statistical information. Within a circle, sectors like 'slices of a pie' represent the quantities involved. The frequency or amount of each quantity is proportional to the angle at the centre of the circle.

bar line chart (KS3) Similar to a bar chart, but for categorical data, the width of bars is reduced so that they appear as lines. The lengths of the bar lines are proportional to the frequencies. Sometimes called bar line graph.

Scatter graph (KS3): A graph on which paired observations are plotted and which may indicate a relationship between the variables. Example: The heights of a number of people could be plotted against their arm span measurements. If height is roughly related to arm span, the points that are plotted will tend to lie along a line.

Line of best fit (KS3): A line drawn on a scatter graph to represent the best estimate of an underlying linear relationship between the variables.

histogram: A particular form of representation of grouped data. Segments along the x-axis are proportional to the class interval. Rectangles are drawn with the line segments as bases. The area of the rectangle is proportional to the frequency in the class. Where the class intervals are not equal, the height of each rectangle is called the frequency density of the class.
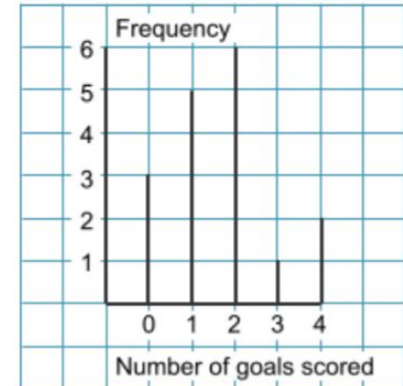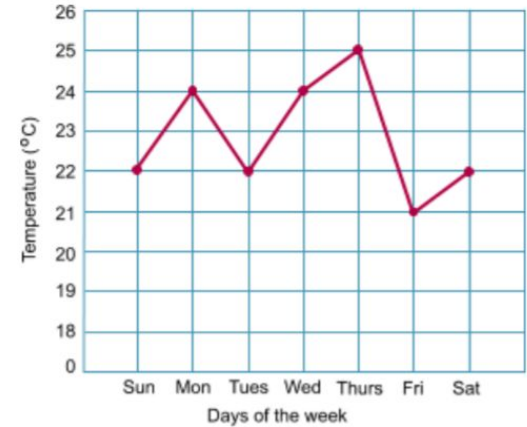
Averages

average (KS2): Loosely an ordinary or typical value, however, a more precise mathematical definition is a measure of central tendency which represents and or summarises in some way a set of data. The term is often used synonymously with 'arithmetic mean', even though there are other measures of average. See median and mode

moving average (KS3): The mean of a set of adjacent observations of fixed size is taken. The mean is calculated for successive sets of the same size to give the moving average. E.g. For example, the moving average of six-month sales may be computed by taking the average of sales from January to June, then the average of sales from February to July, then of March to August, and so on.

mean (KS2): Often used synonymously with average. The mean (sometimes referred to as the arithmetic mean) of a set of discrete data is the sum of quantities divided by the number of quantities. Example: The arithmetic mean of 5, 6, 14, 15 and 45 is (5 + 6 + 14 + 15 + 45) ÷ 5 i.e. 17. More correctly called the arithmetic mean, as there are also other means in mathematics. See mode and median.

arithmetic mean (KS3): The sum of a set of numbers, or quantities, divided by the number of terms in the set. Example: The arithmetic mean of 5, 6, 14, 15 and 45 is (5 + 6 + 14 + 15 + 45) ÷ 5 i.e. 17.

median (KS3): The middle number or value when all values in a set of data are arranged in ascending order. Example: The median of 5, 6, 14, 15 and 45 is 14. When there is an even number of values, the arithmetic mean of the two middle values is calculated. Example: The median of 5, 6, 7, 8, 14 and 45 is (7 + 8) ÷ 2 i.e. 7.5. The median is one example of an average. See also mean, arithmetic mean and mode.

Mode: (KS3): The most commonly occurring value or class with the largest frequency. e.g. the mode of this set of data: 2, 3, 3, 3, 4, 4, 5, 5, 6, 7, 8 is 3 Some sets of data may have more than one mode.

range (KS3) A measure of spread in statistics. The difference between the greatest value and the least value in a set of numerical data.

Statistical language

sample (KS2) A subset of a population. In handling data, a sample of observations may be made from which to draw inferences about a larger population.

probability (KS3) The likelihood of an event happening. Probability is expressed on a scale from 0 to 1. Where an event cannot happen, its probability is 0 and where it is certain its probability is 1. The probability of scoring 1 with a fair dice is 1/6. The denominator of the fraction expresses the total number of equally likely outcomes. The numerator expresses the number of outcomes that represent a 'successful' occurrence. Where events are mutually exclusive and exhaustive the total of their probabilities is 1. probability

scale (KS3) This is a scale between zero and 1, with zero representing the impossibility of an event and 1 representing the fact that the event must happen. The sum of all the probabilities for all the events in a sample space is 1, where the sample space is the set of all possible outcomes of a trial.

correlation (KS3): A measure of the strength of the association between two variables. High correlation implies a close relationship and low correlation a less close one. If an increase in one variable results in an increase in the other, then the correlation is positive. If an increase in one variable results in a decrease in the other, then the correlation is negative. The term zero correlation does not necessarily imply 'no relationship' but merely 'no linear relationship'

proportion (KS2/3) 1. A part to whole comparison. Example: Where £20 is shared between two people in the ratio 3 : 5, the first receives £7.50 which is 3/8 of the whole £20. This is his proportion of the whole. 2. If two variables x and y are related by an equation of the form y = kx, then y is directly proportional to x; it may also be said that y varies directly as x. When y is plotted against x this produces a straight line graph through the origin. 3. If two variables x and y are related by an equation of the form xy = k, or equivalently y = k/x, where k is a constant and x ≠ 0, y ≠ 0 they vary in inverse proportion to each other

proportional reasoning (KS2) Using the mathematics and vocabulary of ratio, proportion and hence fractions and percentages to solve problems.